# Weekly Report

11/02/2015 - 11/08/2015

Li Zongzhuang

Summary

This week I did something.On Monday and Tuesday I was work with my course paper.And the rest days of weekday I read some papers under the command of sennior′s.On the weekend I learned a little of python.

To do list

The next week I will do sme work about the paper of Lu Junhua.Of course we will make a division of labor, and I will do my best.

PS:Iwill explain the MDS in Chinese.

多维尺度(Multidimensionalsealing)是一种传统的寻求保持数据点之差异性(或相似性)的降维方法,它可使得在原数据集合中相近的点仍然靠在一起,远离的点仍然远离。它是非线性的降维方法。

一般而言，MDS 的目的是对一个包含目标的集合，把它的相互接近程度进行可视化表达。

从数学角度来看，MDS 方法就是寻找一组 p 维空间的向量，使得在根据成为 Stress 的标准函数下，各项的欧氏距离矩阵和原矩阵经过某函数所得矩阵尽可能的一致。Stress 函数是用来度量 MDS 表示的图形和原矩阵所给出的各项之间的相互关系的一致性的。我们用 $D_{ij}$ 表示原始空间第 i，j 两点间的距离，$d_{ij}$ 表示对应低维空间中的距离，则这些函数的一般形式如下：

$$\sqrt{\frac{\sum\sum\left(f\left(D_{ij}\right)-d_{ij}\right)^2}{scale}}$$

等式中 $f\left(D_{ij}\right)$ 表示对输入数据变换的函数，scale 是尺度因子，常数，用来保证 Sress 的值在 0 和 1 之间。Sress 值越小，MDS 表现越好。

（计算距离的最常用的方法有欧氏距离算法 (Eucliden distance)和闵氏距离算法(Minkowski distance)。计算公式如下：

$d_{ij}=\sqrt{\sum_a (x_{in}-x_{ja})^2}$ 欧氏距离公式

$d_{ij}=\sqrt[r]{\sum_a |x_{in}-x_{ja}|^r}$ 闵氏距离公式（当 r 等于 2，与欧氏距离相当，r 等于 1，与城区距离相同，r 等于∞，与切比雪夫距离相同。））

常用的 Sress 函数包括 Kruskal Stress，简称“Stress 1”。它的式子为：

$$\sqrt{\frac{\sum\sum\left(f\left(D_{ij}\right)-d_{ij}\right)^2}{\sum\sum {D_{ij}}^2}}$$

$f(D_{ij})$采用什么取决于尺度是度量尺度还是非度量尺度。在度量尺度中，

$f(D_{ij})=D_{ij}$。即原始数据直接与投影之后地图距离进行比较。而在非度量尺度下，

$f(D_{ij})$是一个弱单调函数，它把输入数据变换后的 Stress 变得最小。

从数学角度来看,出现非零的 Stress 值只有一个原因:维数不足"也就是说,对于任意的数据集,把它的结构完美的表现在 2 维或 2 维以卜的空间是不太可能的。当然,并不是 MDS 的 stress 一定为 0 才一能达到使用效果。一定量的失真足可以容忍的。通常，对 MDS 分析结果进行解释要注意两方面，一是各维度意义的解释，二是对各研究对象在空间中的分布特点进行解释。